Generation of pulmonary neuro-endocrine cells and tumors resembling small cell lung cancers from human embryonic stem cells

Huanhuan Joyce Chen[1,*#], Asaf Poran[1,2,#] , Arun M. Unni[1,], Sarah Xuelian Huang[3,@], Olivier Elemento[1,2], Hans-Willem Snoeck[3], Harold Varmus[1,*]

[1] Meyer Cancer Center, Weill Cornell Medicine, New York NY USA. [2] Caryl and Israel Englander Institute for Precision Medicine and Institute for Computational Biomedicine, Weill Cornell Medicine, New York NY USA. [3] Columbia Center for Translational Immunology, Columbia University Medical Center. New York NY USA.

[#] Co-first authors

*To whom correspondence should be addressed

Harold Varmus MD
Meyer Cancer Center
Weill Cornell Medicine
413 East 69th street, BB1322
New York, NY. 10021

Email: varmus@med.cornell.edu

Huanhuan Joyce Chen PhD
Meyer Cancer Center
Weill Cornell Medicine
413 East 69th street, BB1352
New York, NY. 10021

Email: huc2003@med.cornell.edu

@Current address: Center for Stem Cell and Regenerative Medicine, Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center, Houston, Texas, USA.

Human cancers arising from different cell lineages display genotypes with characteristic features, but the lineage-specific factors that influence the differences in genetic profiles have not been identified[1, 2]. Such differences could be explained by models in which a cellular phenotype limits a cell's vulnerability to certain mutations or by models in which observed mutations determine the ultimate cell phenotype. To evaluate the relative merits of these proposals, we are studying the initiation of cancers by genetically modifying cells at discrete stages of differentiation after derivation from human embryonic stem cells (hESCs). We have focused initially on small cell lung cancer (SCLC), the most aggressive type of human lung cancer, characterized by a poor prognosis, the rapid development of resistance to treatment, and nearly universal loss of function of multiple tumor suppressor genes, especially *TP53* and *RB* [3-7]. Earlier studies with mouse models indicate that the normal cell precursors to SCLCs are likely to be pulmonary neuroendocrine cells (PNECs)[8-10]. Building on existing methods for sequentially differentiating hESCs in culture into several types of lung cells, we have now produced PNECs in significant numbers for the first time by blocking signaling through NOTCH receptors (using inhibitors of γ-secretase) and by interfering with expression of the *RB* gene (using inhibitory RNAs) in lung progenitor cells. In contrast, expression of mutated *EGFR* or *KRAS* genes, common in human lung adenocarcinomas that arise from the alveolar cell lineage, has no evident effects on the formation of PNECs. Although PNECs induced by blocking NOTCH and RB signaling do not form xenografts in immune-deficient mice, the cells form subcutaneous tumors resembling early stage human SCLC and bearing neuroendocrine markers if expression of the *TP53* gene is also impaired. These findings imply that PNECs or their committed precursors are vulnerable to oncogenic mutations found specifically in human SCLC, resulting in a significant increase in number of PNECs and conversion to neoplasia. This experimental system provides opportunities to study tumor progression and cancer drug susceptibility and resistance, using human lung cells grown in culture.

Methods have recently been described for generating most of the cell types observed in adult lung tissues from certain lines of hESCs, using combinations of growth factors and chemicals that perturb signaling pathways in a chronological fashion over several weeks (Fig. 1a). Using a protocol developed by Huang et al[11, 12], we have confirmed that by day 3, about 90% of hESCs (the RUES2 and ES02 lines) differentiated into definitive endoderm (DE), triple positive for the markers KIT, EPCAM, and CXCR4 (Extended Data Fig. 1a,b); anterior foregut endoderm (AFE) by day 6; increasing numbers of lung progenitors (LP), SOX2+, NKX2.1+ and FOXA2+, between days 15 and 25 (Extended Data Fig. 1c, d, Extended Data Fig. 3a, b); and then a variety of airway and lung epithelial cells (basal progenitor cells, ciliated cells, goblet cells, club cells, and alveolar type 1 and type 2

cells [13, 14] [AT1 and AT2]) by day 55 (Extended Data Fig. 2a, c). However, this and other protocols produce few PNECs (<0.5%; Fig.1b, c; Extended Data Fig. 2b, c).

Studies of mouse development have suggested that inhibition of signaling via NOTCH receptors might influence cells to adopt a neuroendocrine fate [10, 15-17]. In addition, inactivation of *NOTCH* genes is found in about 25% of SCLCs [5]. We exposed LPs to N-[(3,5-Difluorophenyl)acetyl]-L-alanyl-2-phenyl]glycine-1,1-dimethylethyl ester (DAPT)[18]---a known inhibitor of γ-secretase, the protease that normally cleaves NOTCH receptors to yield a transcriptionally active, mobile, intracellular domain of NOTCH (NICD)[19]---to ask whether the loss of NOTCH signaling might promote the production of PNECs. (We define PNECs here as cells containing a general lung-specific marker, the transcription factor NKX2.1, and expressing one or more genes encoding well-recognized neuroendocrine markers, especially the cell surface-associated protein that includes the calcitonin gene-related peptide (CGRP) [9] or the nuclear transcription factor ASCL1[20, 21].) If exposed to 5 to 10uM DAPT for 30 days, a substantial number of differentiating LP cells (about 8.9 ± 1.9% in RUES2 cells, and 5.6 ± 1.4% in ES02 cells, versus 0.5 ± 0.20% or 0.4 ± 0.2% in control cultures) adopt properties of PNECs, as measured by counting CGRP+ cells with fluorescence-activated cell sorting (FACS) and confirmed qualitatively by immunofluorescence of cells in monolayer cultures with antibodies against NXK2.1 and CGRP (Fig.1b.c, Extended Data Fig. 3 c, d).

We took several approaches to confirm the mechanism by which DAPT induced PNECs. First, we measured the abundance of NICD in extracts from cells treated with DAPT and observed the expected loss of the γ-secretase cleavage product (Fig. 1d). To confirm the expected reduction in NOTCH-mediated signaling, we measured the readout from two NOTCH target genes [22], *HES1* and *HEY1,* and found a marked decrease in levels of transcripts from *HEY1* and a slight but significant decrease of RNA from *HES1* (Extended Data Fig. 4a). We also tested another known inhibitor of γ-secretase, dibenzazepine (DBZ)[23], and produced percentages of PNECs at day 55 similar to those observed with DAPT (Fig. 1e). Finally, we rescued cells from the effects of DAPT by providing NICD, the normal product of γ-secretase-mediated cleavage of NOTCH. Expression of a

tetracycline-inducible transgene encoding NICD from days 25 to 55 prevented the appearance of CGRP+ cells in cultures concurrently exposed to DAPT (Fig. 1 f, g)

To further characterize the presumptive PNECs generated by inhibition of NOTCH signaling, we used high-throughput single-cell RNA sequencing (scRNA-seq) applied to DAPT-treated and untreated cells at day 55 (Fig. 1h-k, Extended Data Fig. 2 b, c, Extended Data Fig. 4 b; see Methods). Clustering of the single-cell profiles revealed one cluster enriched with cells expressing genes that encode CGRP or ASCL1, thus identifying the presumed PNEC cells. In total, the presumed PNEC cells constituted 7.72% of the 9,824 high quality cells (pooled from two biological replicates). Analysis of differential gene expression in the various clusters revealed that cells in cluster 4 exhibit relatively high numbers of transcripts from a set of genes encoding canonical PNEC markers [9, 10] such as *CGRP, ASCL1, GRP, SYP, UCHL1*, and others characteristically expressed in neuroendocrine cells (Fig. 1h, i). In comparison, we detected *CGRP* or *ASCL1* RNA in only 1.3% percent of cells from control cultures not treated with DAPT; scRNA-seq identified only ASCL1 with little CGRP in most of these cells (Extended Data Fig 2c.). Investigation of the other clusters indicates that they express a variety of genes encoding markers specific for several types of lung cells, as also previously noted by Treutlein et al [13] (Extended Data Fig 4b). Some of these marker genes are co-expressed in individual cells, and some are expressed in distinct populations; however, we have not pursued these observations further---for example, to propose differentiation pathways for lung development.

Since a central objective of this work is to assess the influence of known lung cancer genes on the behavior of cells in the lung lineage, we next examined the consequences of expressing or simulating known oncogenic mutations in hESC cultures undergoing differentiation, with or without inhibition of NOTCH signaling (Figs 2 and 3). To that end, we equipped the RUES2 hESC line with Dox-inducible transgenes encoding small hairpin RNAs (shRNAs) that inhibit production of RNA from either of the two tumor suppressor genes most commonly inactivated by mutations in SCLC, the *RB* or *TP53* genes (Extended Data Fig. 5a, b, and Fig. 3 a), or with inducible transgenes encoding oncogenes commonly encountered in lung adenocarcinomas, mutant *EGFR* or mutant *KRAS* [24] (Fig. 3 b).

Induction of *RB*-specific shRNA in RUES2 cells differentiating to form LCs between days 25 and 50 blocked the production of RB protein (Extended Data Fig. 5a) but not the production of the closely related proteins p107 and p130 (Extended Data Fig. 5b). Reduced expression of the *RB* gene was associated with a significantly increased number of CGRP+ NKX2.1+ cells (putative PNECs) from 7.5 ± 2.0% to 37.8 ± 8.2% (Fig. 2 a-d), as measured by FACS, but only when cells were also exposed to DAPT to inhibit processing of NOTCH (Fig.2d). Similarly, the proportions of cells expressing the PNEC transcription factor ASCL1 and associated markers NCAM1 and CHGA were also significantly increased (Fig. 2e). In contrast, induction of *TP53*-specific shRNA during the same interval had no effect on the number of CGRP+ cells, with or without DAPT and with or without induction of *RB*-shRNA (Fig. 3c-e). Similarly, induction of mutant forms of EGFR and KRAS proteins between days 25 and 55 (Fig. 3b) failed to increase the number of CGRP+ cells grown in the absence or presence of DAPT (Fig. 3f, g).

To examine the transcriptional phenotypes of cultures of differentiated (day 55) RUES2 cells in which both NOTCH and RB signaling were inhibited, we turned again to scRNA-seq. Similar clustering and differential expression analyses indicate the presence of multiple cell populations in our cultures, including an expanded PNEC-like cell compartment (11.7%), expressing markers similar to the PNEC-like day 55 cells in which RB was not reduced (Extended Data Fig. 6e). (We attribute the relatively modest increase in PNEC-like cells as judged by scRNA-seq, compared to the increase measured by FACS, to differences in the sensitivity of the detection methods.) In addition, cells in other clusters contained RNA derived from a variety of genes encoding markers specific for several types of lung cells, as also observed using scRNA-seq to examine cultures treated only with DAPT (Extended Data Fig 7a).

To further characterize the effects of the reduction of RB on the transcriptional profile of the PNEC cells, we performed a differential expression analysis that compared PNECs appearing after DAPT treatment with PNECs appearing after treatment with both DAPT and Dox (Extended Data Fig 8a). Subsequent gene function enrichment analysis [25]

indicated that the most differentially expressed genes following reduction of RB gene expression are involved in regulation of nerve development, apoptosis, TP53 signal transduction, and other processes (Extended Data Fig 8b). When investigating the heterogeneity within the DAPT-induced PNEC cell cluster, we detected three sub-populations (Extended Data Fig 8a) with substantially different expression profiles (Extended Data Fig 8c). A similar analysis of PNECs from DAPT-treated cultures in which RB levels were reduced also revealed three subpopulations, but with transcriptional profiles different from those in cultures in which RB expression was not altered (Extended Data Fig 8 a, c).

To ask whether inhibition of NOTCH signaling, coupled with a reduction of RB protein, produces a transcriptional program that resembles the program in human SCLC, we compared the scRNA-seq profiles of PNECs and non-PNECs from day 55 RUES2 cells, with normal or reduced levels of *RB* gene expression, to the published RNA profiles from 29 early-stage (stage Ia or Ib) human SCLCs [5, 6] (Extended Data Fig 7 b). This analysis confirmed that PNEC expression profiles more closely resemble SCLC profiles than do the profiles from non-PNECs. In addition, the similarity to SCLC profiles is greater ($P < 2.2e-16$ by two sided Kolmogorov–Smirnov test) in PNECs in which RB levels were reduced than in PNECs in which RB levels were not perturbed.

To assess the ability of cells in differentiated RUES2 cultures to form tumors, we performed subcutaneous injections of day 55 cells treated in various ways into immune-deficient NSG mice (NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ)[26]. The four tested cell populations all contained presumptive PNECs (CGRP+ by FACS), ranging from about 10 to about 40 percent of the cultures. No growths greater than 250 mm in diameter were observed with parental cells exposed to DAPT alone or with cells carrying the *TP53*-shRNA or the *RB*-shRNA expression cassettes and treated with Dox (Fig 4 a). In contrast, cells with both shRNA cassettes and treated with DAPT and Dox formed tumors at 14 of 19 injection sites. In general, these tumors were about 1 cm in diameter, formed of compact, darkly staining cells, morphologically resembling SCLC in mice and humans (Fig 4 b), and not locally invasive. The origin of the tumor cells was confirmed by detection of GFP encoded

by a component of the *RB*-shRNA cassette in the RUES2 cells nuclei (Extended Data Fig. 9 b) as well as positive staining of human nuclei (Extended Data Fig. 9 c). A PNEC-like phenotype was documented using IHC to display the neuroendocrine biomarkers CGRP, NCAM1, and ASCL1, as well as the lung marker NKX2.1(Figure 4b). We also excluded the possibility of confusion with the teratomas known to be formed in mice injected with hESCs. Teratomas formed in NSD mice with undifferentiated RUES2 cells contained embryonic tissue markers[27], such as α-fetoprotein, Nanog, Oct4, and SSEA4 (Extended Data Fig. 9 d), and exhibited morphological features different from tumors formed with our differentiated cultures in which NOTCH, RB, and P53 pathways were disrupted (Extended Data Fig. 9a).

In summary, we have developed a human cell-based model in tissue culture for the initiation of early stage tumors with several properties of SCLC. To do this, we used an existing protocol to differentiate hESCs, induced PNEC-like cells by inhibition of NOTCH signaling, expanded the number of PNEC cells by blocking the expression of the *RB* tumor suppressor gene, and conferred a tumorigenic phenotype on those cells with an additional block to expression of *TP53*, another tumor suppressor gene commonly inactivated in SCLC. Since the subcutaneous tumors grown from these cells in immunodeficient mice appear to have low tumor potency (slow-growing and non-invasive), it is likely that this system will allow studies of tumor progression. In addition, it should be possible to examine cells at different stages of tumor development for susceptibility and resistance to therapeutic strategies, in a manner similar to that used in an earlier study, in which a rare form of glioma was derived from hESCs by genetic manipulation of neural precursors [28]. Since we have been able to generate early stage SCLC-like growths only after reducing expression of tumor suppressor genes implicated in human SCLC and inducing neuroendocrine cell types, but not with oncogenes known to be drivers of other types of human lung cancer, it seems likely that differentiation creates cell types, like LPs and PNECs, that are vulnerable to specific types of oncogenic lesions, such as inactivation of certain tumor suppressor genes. This may help to account for the widely recognized correlations between cancer genotypes and lineage phenotypes in human neoplasias.

Study approval

All embryonic stem cell studies were approved by the Tri-Institutional ESCRO committee (Weill Cornell Medicine, Sloan-Kettering Cancer Center and Rockefeller University). All animal protocols in this study were approved by the IACUC committees of Weill Cornell Medicine, Cornell University.

Accession ID

Single cell sequence data will be deposited in NCBI SRA.

Author Contributions

HJC and AP performed the experiments. All authors participated in the design and interpretation of some or all experiments. HV, HJC, and AP wrote the manuscript, and all authors suggested editorial changes. HV conceived the study and recruited the collaborating partners.

Acknowledgements

We thank Oksana Mashadova and Sukanya Goswami in the Varmus Laboratory for technical support; Dennis Fei and John Ferrarone in the Varmus Laboratory, Yawen Chen in the Snoeck Laboratory, Rahul Satija (New York Genome Center), Viviane Tabar (Memorial Sloan Kettering Cancer Center) and Shuibing Chen (Weill Cornell Medicine) for useful advice. Supported by an award (to H.V. and O.E.) from the US Department of Defense (LC160136), funds from the Meyer Cancer Center, Weill Cornell Medicine (to H.V.), an Arnold O. Beckman Postdoctoral fellowship (to H.J.C), and a Weill Cornell graduate fellowship (to A.P).

Competing Financial Interests The authors have no financial conflicts of interest.

ABBREVIATIONS:

PNECS: pulmonary neuro-endocrine cells; **LUSD:** lung squamous cell carcinoma; **LUAD:** lung adenocarcinoma; **SCLC:** small cell lung cancer; **hESCs:** human embryonic stem cells**; SOX2:** SRY-Box 2; **SPB:** pulmonary-associated surfactant protein B; **SPC:** pulmonary-associated surfactant protein C; **CC10:** The amino acid sequence of Clara Cell 10 kDa secretory protein; **AT2:** alveolar type 2 cells; **EGFR:** Epidermal Growth Factor Receptor; **P63:** transformation-related protein 63; **FOXA2:** Forkhead Box A2; **shRNA:** short hairpin RNA; **HES1:** hairy and enhancer of split-1; **HEY1:** Hes Related Family BHLH Transcription Factor With YRPW Motif 1; **HES6:** hairy and enhancer of split-6; **NE:** neuroendocrine; **NKX2.1:** thyroid transcription factor-1 (TTF-1), also known as Nkx2.1; **NOTCH:** Notch homolog 1; **KRAS:** V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog; **EGFR:** Epidermal growth factor receptor; **TP53:** tumor protein P53; **RB:** retinoblastoma protein; **PTEN:** Phosphatase and tensin homolog; **FACS:** Fluorescence-activated cell sorting; **cDNA:** complementary DNA; **DAPT:** a γ-secretase inhibitor, blocking NOTCH signaling via preventing proteolytic cleavage of the Notch intracellular domain **(NICD)**. **CGRP:** calcitonin gene-related peptide, symbols: CALCA; **GRP:** gastrin-releasing peptide; **SYP:** Synaptophysin; **ASCL1:** Achaete-Scute Family BHLH Transcription Factor 1; **PGP9.5:** Protein Gene Product 9.5., also known as ubiquitin C-terminal hydrolase 1. (UCHL-1); **NCAM1:** Neural Cell Adhesion Molecule 1; **CHGA:** Chromogranin A or parathyroid secretory protein 1; **CHGB:** Chromogranin B; **Oct4:** octamer-binding transcription factor 4; **NANOG:** homeobox transcription factor; **SSEA4:** stage-specific embryonic antigen-4; **RNAseq:** RNA sequencing; **tSNE:** t-Distributed Stochastic Neighbor Embedding analysis; **DBZ:** Diazepine inhibitor of γ-secretase, also known as YO-01027; **FACS:** fluorescence activated cell sorting; **Drop-seq:** a technique analyzing mRNA transcripts from droplets of individual cells in a highly parallel fashion; **10X:** 10X Genomics single-cell 3'mRNA-sequence; **RUES2:** Rockefeller University Embryonic Stem Cell Line 2; **ES02:** HES2 human embryonic stem cell line (NIH codes:ES02).

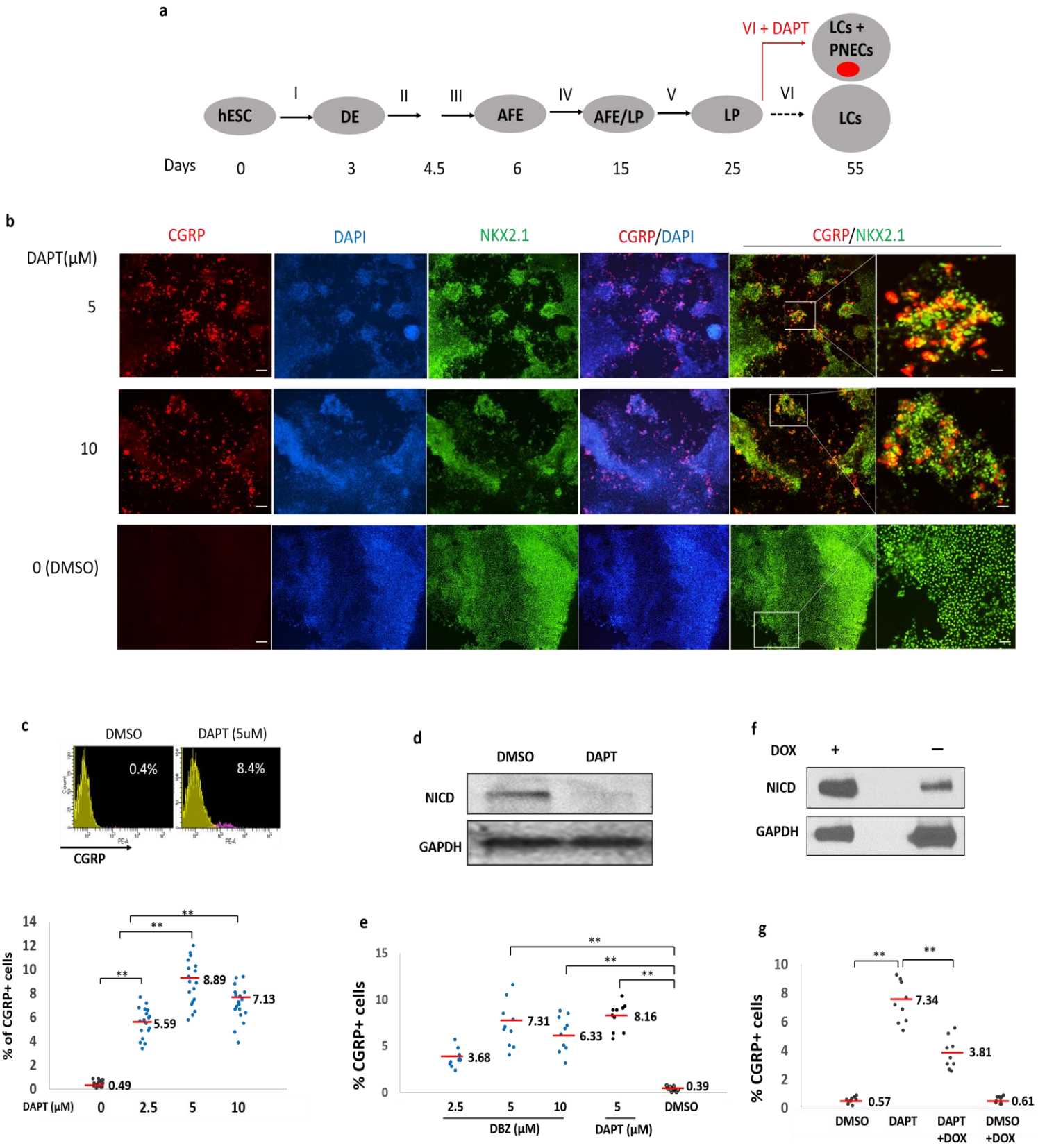
Growth factors & Hormones--- FGF-2: fibroblast growth factor 2; **N2:** a chemically defined, serum-free supplement based on Bottensteins N-1 formulation; **B27:** neural cell culture supplements; **Glutamax:** a stabilized form of L-glutamine; **Y-27632:** an inhibitor of Rho-associated, coiled-coil containing protein kinase (ROCK); **BMP4:** Bone Morphogenetic Protein 4; **bFGF:** *Basic fibroblast growth factor;* **Activin A:** *a member of the transforming growth factor beta (TGF-β) family;* **SB431542:** a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7; **IWP2:** an inactivator of Porcn function and inhibitor of Wnt production; **CHIR99021:** inhibitor of glycogen synthase kinase 3 (GSK-3) and WNT signaling agonist; **FGF10:** Fibroblast growth factor 10; **FGF7(KGF):** Fibroblast Growth Factor 7 (Keratinocyte Growth Factor); **ATRA:** all-*trans* retinoic acid; **DCI:** mixture of Dexamethasone, 8-bromo-cAMP and IBMX (3,7-dihydro-1-methyl-3-(2-methylpropyl)-1*H*-purine-2,6-dione).

REFERENCES

1. Garraway, L.A. & Lander, E.S. Lessons from the cancer genome. *Cell* **153**, 17-37 (2013).
2. Cancer Genome Atlas Research, N. et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nature genetics* **45**, 1113-1120 (2013).
3. Pietanza, M.C., Byers, L.A., Minna, J.D. & Rudin, C.M. Small cell lung cancer: will recent progress lead to improved outcomes? *Journal of the American Association for Cancer Research* **21**, 2244-2255 (2015).
4. Gazdar, A.F., Bunn, P.A. & Minna, J.D. Small-cell lung cancer: what we know, what we need to know and the path forward. *Nature reviews. Cancer* **17**, 765 (2017).
5. George, J. et al. Comprehensive genomic profiles of small cell lung cancer. *Nature* **524**, 47-53 (2015).
6. Peifer, M. et al. Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nature genetics* **44**, 1104-1110 (2012).
7. Semenova, E.A., Nagel, R. & Berns, A. Origins, genetic landscape, and emerging therapies of small cell lung cancer. *Genes & development* **29**, 1447-1462 (2015).
8. Sutherland, K.D. et al. Cell of origin of small cell lung cancer: inactivation of Trp53 and Rb1 in distinct cell types of adult mouse lung. *Cancer cell* **19**, 754-764 (2011).
9. Song, H. et al. Functional characterization of pulmonary neuroendocrine cells in lung development, injury, and tumorigenesis. *Proc Natl Acad Sci U S A* **109**, 17531-17536 (2012).
10. Linnoila, R.I. Functional facets of the pulmonary neuroendocrine system. *Laboratory investigation* **86**, 425-444 (2006).
11. Huang, S.X. et al. The in vitro generation of lung and airway progenitor cells from human pluripotent stem cells. *Nature protocols* **10**, 413-425 (2015).
12. Huang, S.X. et al. Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. *Nature biotechnology* **32**, 84-91 (2014).
13. Treutlein, B. et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509**, 371-375 (2014).
14. Warburton, D., Wuenschell, C., Flores-Delgado, G. & Anderson, K. Commitment and differentiation of lung cell lineages. *Biochemistry and cell biology* **76**, 971-995 (1998).
15. Ito, T. et al. Basic helix-loop-helix transcription factors regulate the neuroendocrine differentiation of fetal mouse pulmonary epithelium. *Development* **127**, 3913-3921 (2000).
16. Morimoto, M., Nishinakamura, R., Saga, Y. & Kopan, R. Different assemblies of Notch receptors coordinate the distribution of the major bronchial Clara, ciliated and neuroendocrine cells. *Development* **139**, 4365-4373 (2012).

17. Shan, L., Aster, J.C., Sklar, J. & Sunday, M.E. Notch-1 regulates pulmonary neuroendocrine cell differentiation in cell lines and in transgenic mice. *Am J Physiol Lung Cell Mol Physiol.* **292**, L500-509 (2007).
18. Geling, A., Steiner, H., Willem, M., Bally-Cuif, L. & Haass, C. A gamma-secretase inhibitor blocks Notch signaling in vivo and causes a severe neurogenic phenotype in zebrafish. *EMBO reports* **3**, 688-694 (2002).
19. Schroeter, E.H., Kisslinger, J.A. & Kopan, R. Notch-1 signalling requires ligand-induced proteolytic release of intracellular domain. *Nature* **393**, 382-386 (1998).
20. Borges, M. et al. An achaete-scute homologue essential for neuroendocrine differentiation in the lung. *Nature* **386**, 852-855 (1997).
21. Borromeo, M.D. et al. ASCL1 and NEUROD1 Reveal Heterogeneity in Pulmonary Neuroendocrine Tumors and Regulate Distinct Genetic Programs. *Cell reports* **16**, 1259-1272 (2016).
22. Iso, T., Kedes, L. & Hamamori, Y. HES and HERP families: multiple effectors of the Notch signaling pathway. *Journal of cellular physiology* **194**, 237-255 (2003).
23. Milano, J. et al. Modulation of notch processing by gamma-secretase inhibitors causes intestinal goblet cell metaplasia and induction of genes known to specify gut secretory lineage differentiation. *Toxicological sciences* **82**, 341-358 (2004).
24. Cancer Genome Atlas Research, N. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543-550 (2014).
25. Chen, J., Bardes, E.E., Aronow, B.J. & Jegga, A.G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic acids research* **37**, W305-311 (2009).
26. Shultz, L.D. et al. Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. *Journal of immunology* **174**, 6477-6489 (2005).
27. Liu, A. et al. Diagnostic utility of novel stem cell markers SALL4, OCT4, NANOG, SOX2, UTF1, and TCL1 in primary mediastinal germ cell tumors. *The American journal of surgical pathology* **34**, 697-706 (2010).
28. Funato, K., Major, T., Lewis, P.W., Allis, C.D. & Tabar, V. Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation. *Science* **346**, 1529-1533 (2014).

Figure 1.



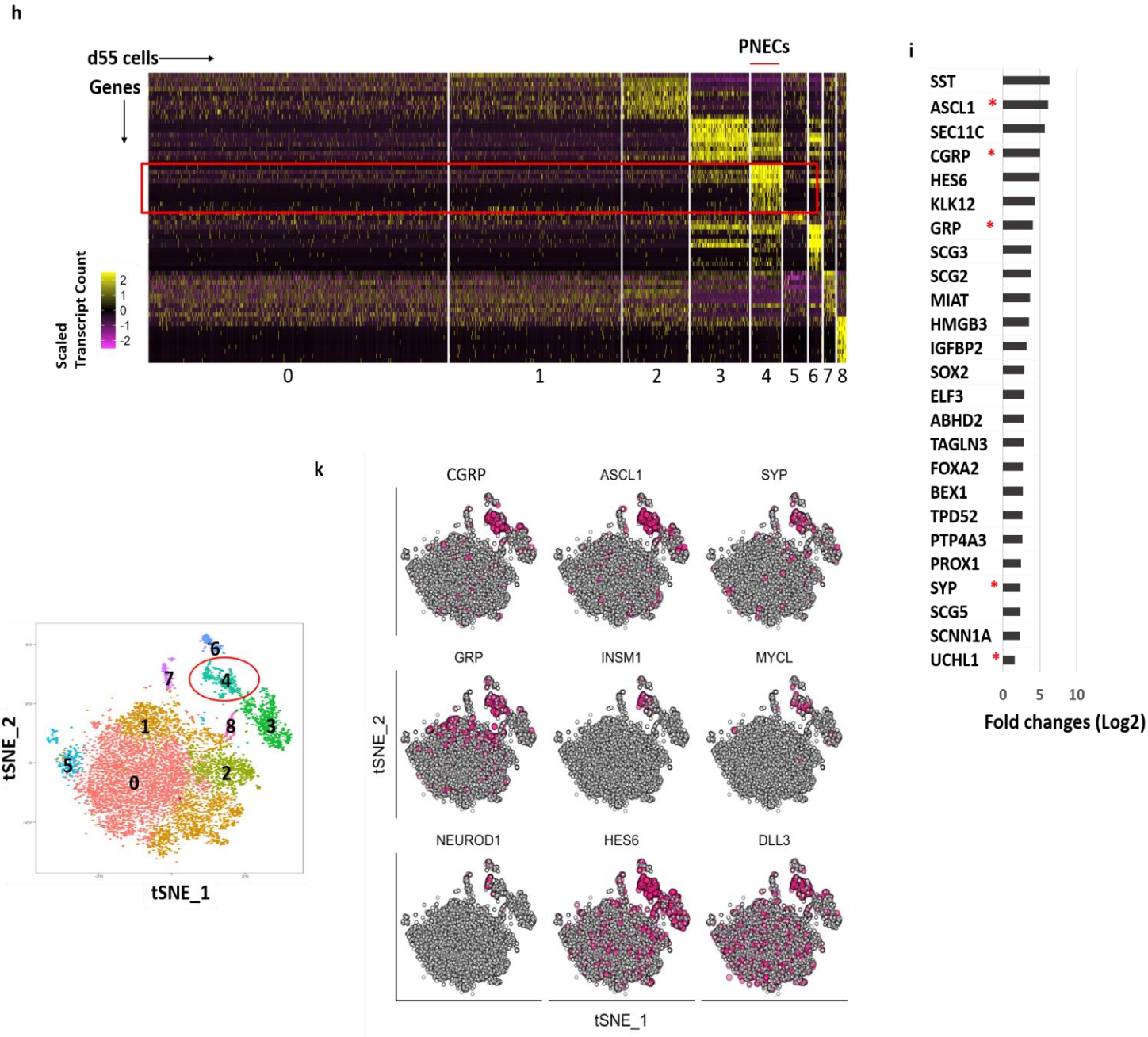
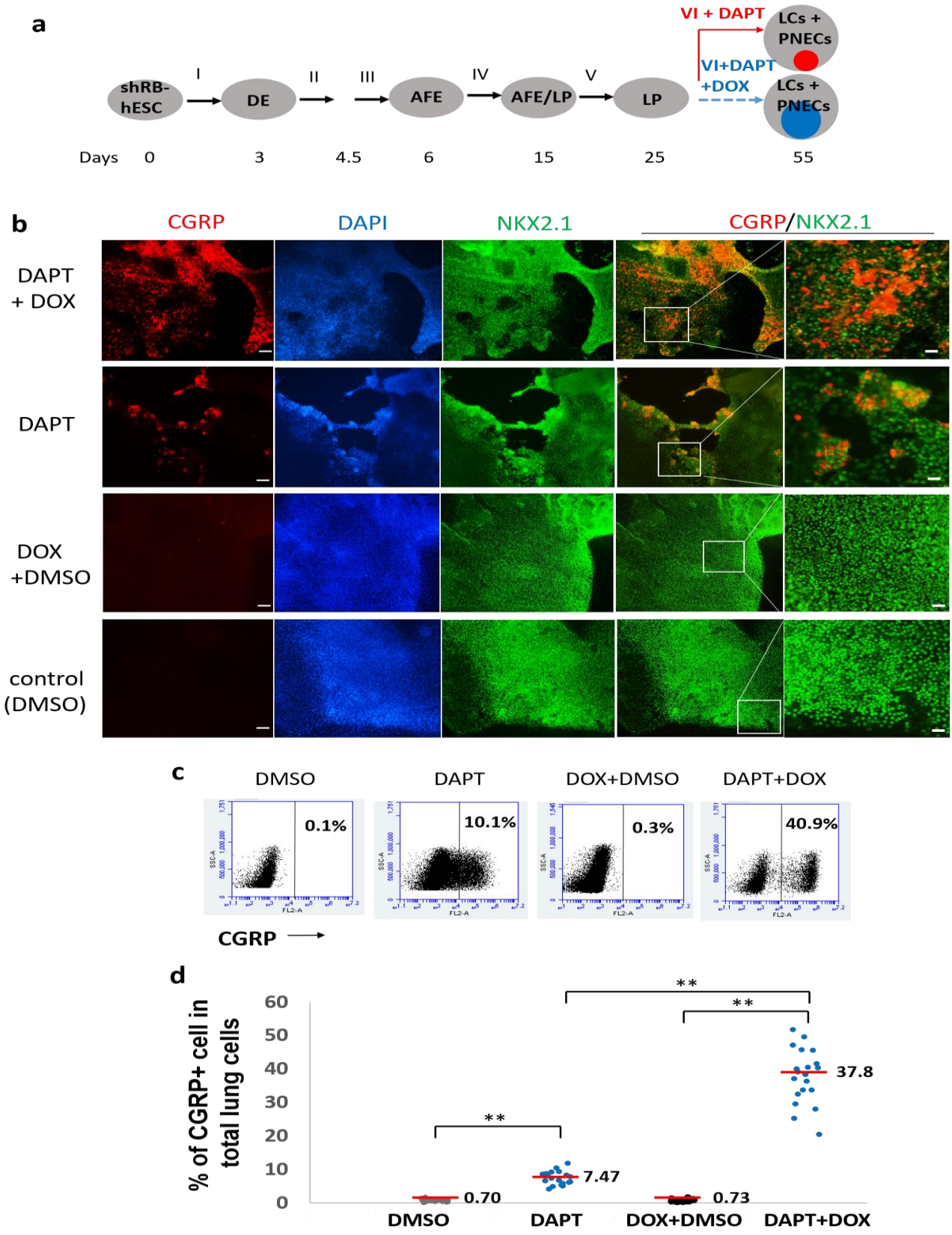


Figure 1. Generation and characterization of PNECs through directed differentiation of hESCs and suppression of NOTCH. a. Schematic of the protocol used to generate PNECs by stepwise differentiation of human embryonic stem cells (hESCs) to form definitive endoderm (DE) by day 3, anterior foregut endoderm (AFE) by day 6, and increasing numbers of lung progenitor cells (LPs) from days 15 to 25, using the differentiation mixtures I to V (defined in Methods; see Extended Data Figs 1 and 3). LPs were further differentiated in mixture VI from days 25 to 55 into the major types of lung cells (LCs) found in mature human lung parenchyma and airway epithelium [13,14]. Addition of DAPT to mixture VI induced formation of pulmonary neuroendocrine cells (PNECs; red dot), as described in the text. **b.** Detection of putative PNECs by IHC after treatment with DAPT. ESCs from the RUES2 line were differentiated according to the protocol in panel **a** to day 55 then stained to detect CGRP, NKX2.1, or both, with the indicated antisera; nuclei were detected by staining with DAPI. Scale bars: 100 μM (left) and 20 μM (right)**. c.** Percentages of CGRP+ cells were determined at day 55 by FACS and displayed as flowcytometry data (red, CGRP+; yellow, CGRP-) and a scatter graph (below). **d-g.** Confirmation of mechanism of action of DAPT as inhibitor of Y-secretase cleavage of NOTCH. **d.** DAPT (5uM) treatment from day 25 to day 55 decreased level of NOTCH intracytoplasmic domain (NICD) in day 55 lung cells, as detected by western blot. **e.** LPs treated with another Y-secretase inhibitor, DBZ from day25 to 55, also form CGRP+ cells. **f-g.** Constitutive expression of NICD prevents the appearance of CGRP+ cells co-treated with DAPT. RUES2 cells carrying a doxycycline-inducible NICD were differentiated to form LPs and then treated with DOX, DAPT, or both for 30 days. Panel **f** demonstrates the induction of NICD by DOX with Western blot; panel **g** shows by FACS that DOX (to induce expression of NICD) inhibits DAPT-mediated appearance of CGRP+ cells. ** $P < 0.01$, * $P < 0.05$ by ANOVA test or (for panel **c**) by Student *t* test. Horizontal red lines denote average values; number of biological repeats (n) = 18 for panel **c**; n=10 for panel **e**; n=9 for panel **g**. **h-k.** Single cell RNA profiling of day 55 lung cells derived from RUES2 cells treated with DAPT (5μM) from day 25 to day 55. **h**. Heatmap representing scaled expression of the most differentially expressed genes specific to different cell clusters. Rows represent genes and columns represent cells **i.** Putative PNEC markers differentially expressed in the PNEC-like cell cluster number 4 in panel **h**. Bars indicate log fold-change versus non-PNEC cells. Asterisks indicate canonical PNEC markers. (See full gene list in Extended Data File 1.) **j.** Reduced-dimensionality t-Distributed Stochastic Neighbor Embedding (tSNE) map colored by cluster assignment (see methods). **k.** Individual cells positive for PNECs markers and other genes associated with neuroendocrine differentiation are denoted by red dots.

Figure 2.



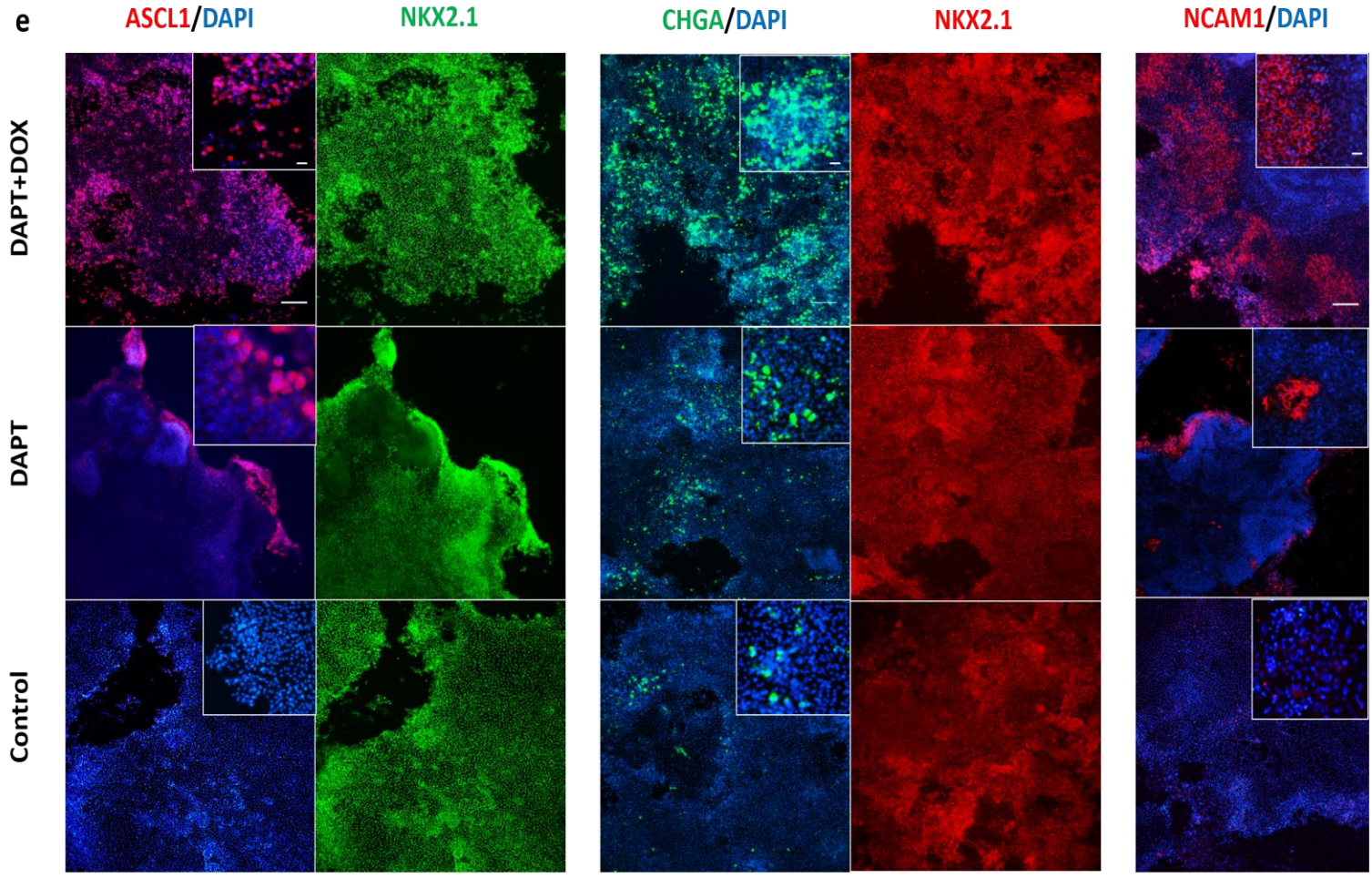


Figure 2. Inhibited expression of the RB tumor suppressor gene augments induction of PNECs by DAPT. a. Schematic of PNEC production from hESC cells carrying a tetracycline-inducible transgene that produces shRNA targeting RB1 (shRB). The format is similar to **Fig.1a,** except that mixture VI is supplemented with DAPT (5μM) with or without doxycycline (DOX). Internal colored circles at day 55 indicate PNECs induced by DAPT (red) or by DAPT and DOX (blue). **b-d.** Increased numbers of putative PNECs detected by co-staining for CGRP and NKX2.1 (panel **b**) or by FACS sorting with anti-human CGRP antibody (panels **c** and **d**) as described in the legend to Fig. 1 and Methods. ** $P < 0.01$ by ANOVA test; in panel **d**, Horizontal red lines denote average values; n= 20; Scale bars: 100 μM (left) and 20 μM (right)**. e.** Expression of shRNA-RB increases the percentages of ASCL1+, CHGA+ or NCAM+ cells. The indicated markers were detected by immunostaining as in panel **b.** Scale bars: 100 μM and 20 μM (in small window)

Figure 3.

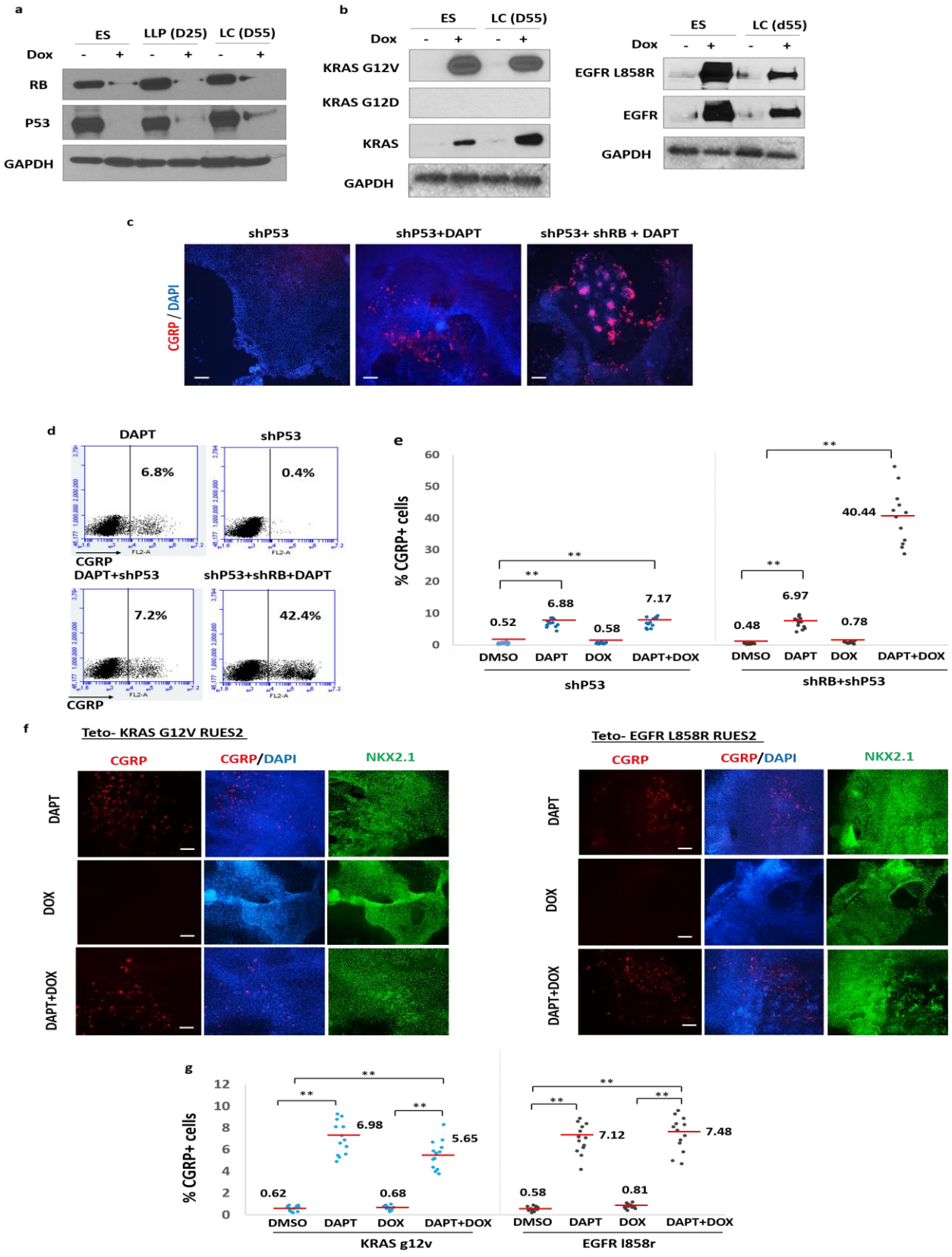
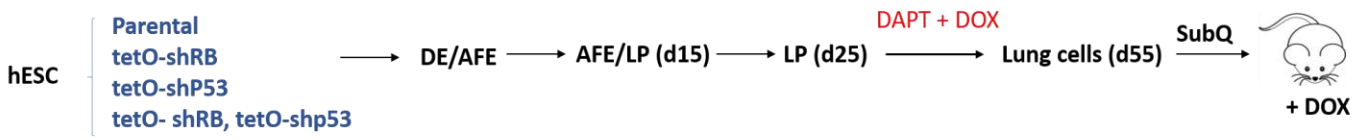


Figure 3. Inhibition of the p53 tumor suppressor gene or expression of two common lung cancer oncogenes do not augment production of PNECs in the presence and absence of DAPT. RUES2 cells carrying DOX-inducible transgenes that encode shRNAs targeting *P53* or *RB* or mutant alleles of *KRAS* (G12V) or *EGFR* (L858R) were used as in the experimental protocol shown in **Fig.2a** to measure the effects of DOX-induction of the indicated transgenes from day 25 to day 55 on production of PNECs. **a.** DOX-dependent expression of oncogenes or shRNA's in RUES2 cells. Western blots show RB and P53 tumor suppressor proteins after DOX-induction of shRNA cassettes. **b.** Production of proteins encoded by DOX-regulated transgenes encoding KRAS-G12V (left panel) and EGFR-L858R (right panel). Proteins were detected by the indicated antibodies (see Methods). **c-g.** Decreased P53 or production of mutant KRAS or EGFR proteins fail to increase the percentage of PNECs. The indicated RUES2 cell lines were tested for the appearance of PNECs at day 55 by immunofluorescence staining for CGRP (panels **c** and **f**) and by FACS (panels **d, e, g**). See text for interpretation. Scale bars: 100 μM in panels **c;** 200 μM in panels **f**, ** $P < 0.01$ by ANOVA test; horizontal red lines denote average values; n= 12 for **e**, and 13 for **g**.

Figure 4.

a



Tumor formation in immunodeficient mice

Cells	% CGRP+ cells in total injected cells	Tumors / injection (≥ 250mm^3)
Parental (DAPT alone)	7.6 ± 1.0	0/12
DAPT+shRB	39.6 ± 4.4	0/14
DAPT+shP53	8.0 ± 1.3	0/11
DAPT+shRB+shP53	41.9 ± 4.6	14/19 **

b

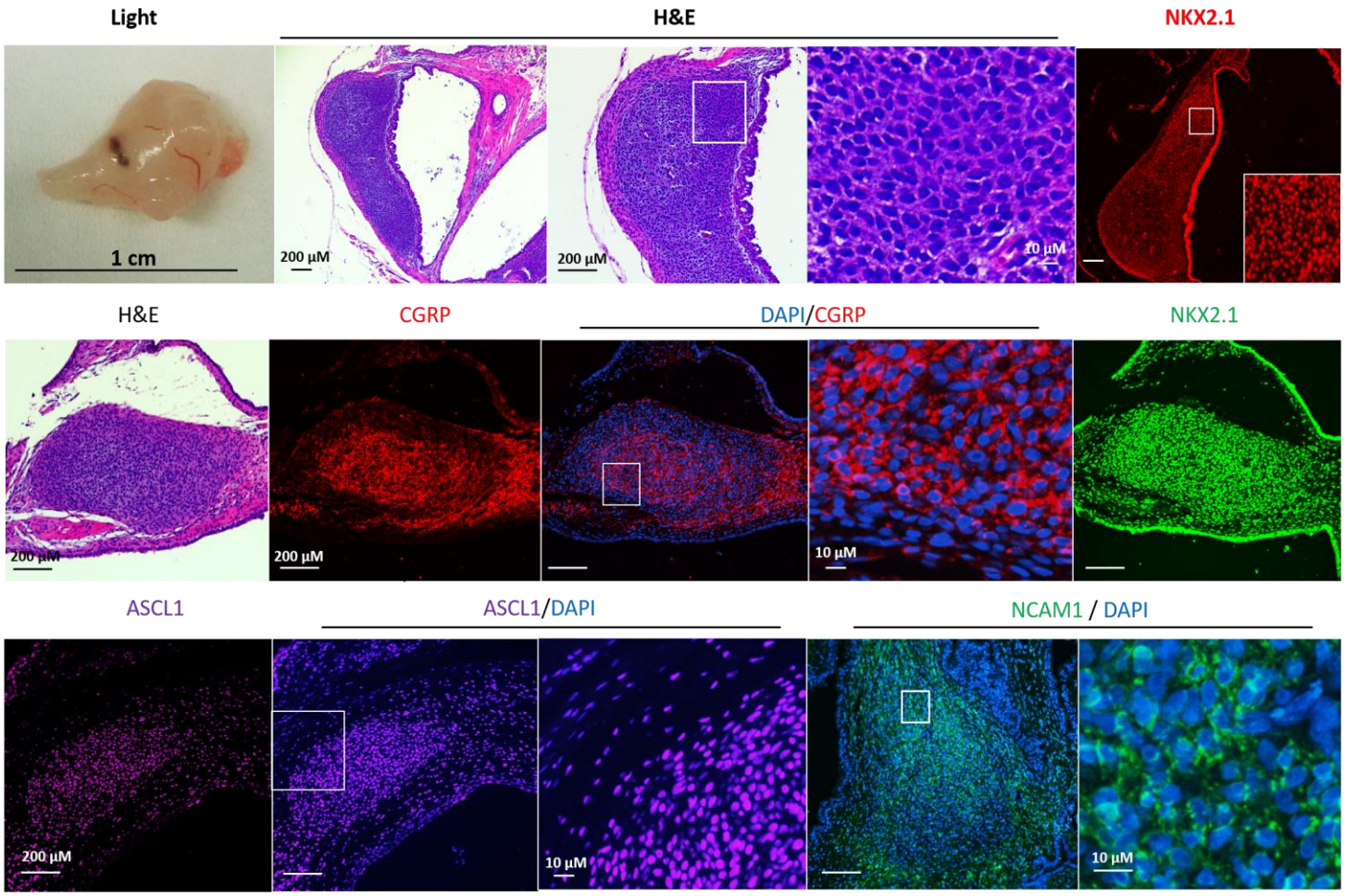


Figure 4. Characterization of xenografts formed with hESC – derived lung cells. a. Schematic representation of tumorigenesis experiment. The indicated transgenic and control lines of RUES2 hESCs were differentiated and grown in DAPT and Dox from d25 to d55. At day 55, total lung cells were injected subcutaneously into NSG mice. Xenografts grew into visible tumors after 6-7 weeks only from cells containing transgenes encoding shRNAs for both RB and P53. ** $P < 0.01$ by Fisher test. **b.** Characterization of xenografts. **Upper panel:** left segment, a representative tumor by light microscopy; middle three segments, H & E staining of that tumor at different magnifications; right segment, staining for NKX2.1; **Middle and Lower panels:** Additional samples stained with H & E, DAPI, or antibodies specific for the indicated marker proteins.